2.8 The Invertible Matrix Theorem I

In Section 2.6, we defined an \(n \times n\) invertible matrix \(A\) to be a matrix such that there exists an \(n \times n\) matrix \(B\) satisfying \(AB = BA = I_n\). There are, however, many other important and useful viewpoints on invertibility of matrices. Some of these we have already encountered in the preceding two sections, while others await us in later chapters. It is worthwhile to begin collecting a list of conditions on an \(n \times n\) matrix \(A\) that are

(a) The inverse of an invertible upper triangular matrix is upper triangular. Repeat for an invertible lower triangular matrix.

(b) The inverse of a unit upper triangular matrix is unit upper triangular. Repeat for a unit lower triangular matrix.

29. In this problem, we prove that the LU decomposition of an invertible \(n \times n\) matrix is unique in the sense that, if \(A = L_1U_1\) and \(A = L_2U_2\), where \(L_1, L_2\) are unit lower triangular matrices and \(U_1, U_2\) are upper triangular matrices, then \(L_1 = L_2\) and \(U_1 = U_2\).

(a) Apply Corollary 2.6.12 to conclude that \(L_2\) and \(U_2\) are invertible, and then use the fact that \(L_1U_1L_1^{-1}L_2^{-1} = U_2^{-1}\).

(b) Use the result from (a) together with Theorem 2.2.22 and Corollary 2.2.23 to prove that \(L_1^{-1}L_2 = I_n\) and \(U_1^{-1}U_2 = I_n\), from which the required result follows.

30. QR Factorization: It can be shown that any invertible \(n \times n\) matrix has a factorization of the form

\[A = QR\]

where \(Q\) and \(R\) are invertible, \(R\) is upper triangular, and \(Q\) satisfies \(Q^TQ = I\) (i.e., \(Q\) is orthogonal).

Determine an algorithm for solving the linear system \(Ax = b\) using this QR factorization.

For Problems 31–33, use some form of technology to determine the LU factorization of the given matrix. Verify the factorization by computing the product \(LU\).

31. \(A = \begin{bmatrix} 3 & -2 \\ 2 & 7 \\ -5 & 11 \end{bmatrix}\)

32. \(A = \begin{bmatrix} 27 & -19 & 32 \\ 15 & -16 & 9 \\ 23 & -13 & 51 \end{bmatrix}\)

33. \(A = \begin{bmatrix} 34 & 13 & 19 & 22 \\ 53 & 17 & -71 & 20 \\ 21 & 37 & 63 & 59 \\ 81 & 93 & -47 & 39 \end{bmatrix}\)
mathematically equivalent to its invertibility. We refer to this theorem as the Invertible Matrix Theorem. As we have indicated, this result is somewhat a “work in progress,” and we shall return to it later in Sections 3.2 and 4.10.

**Theorem 2.8.1** (Invertible Matrix Theorem)

Let $A$ be an $n \times n$ matrix with real elements. The following conditions on $A$ are equivalent:

(a) $A$ is invertible.

(b) The equation $Ax = b$ has a unique solution for every $b$ in $\mathbb{R}^n$.

(c) The equation $Ax = 0$ has only the trivial solution $x = 0$.

(d) $\text{rank}(A) = n$.

(e) $A$ can be expressed as a product of elementary matrices.

(f) $A$ is row-equivalent to $I_n$.

**Proof** The equivalence of (a), (b), and (d) has already been established in Section 2.6 in Theorems 2.6.4 and 2.6.5, as well as in Corollary 2.6.6. Moreover, the equivalence of (a) and (e) was already established in Theorem 2.7.5.

Next we establish that (c) is an equivalent statement by proving that (b) $\implies$ (c) $\implies$ (d). Assuming that (b) holds, we can conclude that the linear system $Ax = 0$ has a unique solution. However, one solution is evidently $x = 0$, hence this is the unique solution to $Ax = 0$, which establishes (c). Next, assume that (c) holds. The fact that $Ax = 0$ has only the trivial solution means that, in reducing $A$ to row-echelon form, we find no free parameters. Thus, every column (and hence every row) of $A$ contains a pivot, which means that the row-echelon form of $A$ has $n$ nonzero rows; that is, $\text{rank}(A) = n$, which is (d).

Finally, we prove that (a) $\implies$ (f) $\implies$ (a). If (a) holds, we can left multiply $I_n$ by a product of elementary matrices (corresponding to a sequence of elementary row operations applied to $I_n$) to obtain $A$. This means that $A$ is row-equivalent to $I_n$, which is (f). Last, if $A$ is row-equivalent to $I_n$, we can write $A$ as a product of elementary matrices, each of which is invertible. Since a product of invertible matrices is invertible (by Corollary 2.6.10), we conclude that $A$ is invertible, as needed.

**Exercises for 2.8**

**Skills**

- Know the list of characterizations of invertible matrices given in the Invertible Matrix Theorem.
- Be able to use the Invertible Matrix Theorem to draw conclusions related to the invertibility of a matrix.

**True-False Review**

For Questions 1–4, decide if the given statement is true or false, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. If the linear system $Ax = 0$ has a nontrivial solution, then $A$ can be expressed as a product of elementary matrices.

2. A $4 \times 4$ matrix $A$ with rank($A$) = 4 is row-equivalent to $I_4$.

3. If $A$ is a $3 \times 3$ matrix with rank($A$) = 2, then the linear system $Ax = b$ must have infinitely many solutions.

4. Any $n \times n$ upper triangular matrix is row-equivalent to $I_n$. 
2.9 Chapter Review

Problems

1. Use part (c) of the Invertible Matrix Theorem to prove that if \( A \) is an invertible matrix and \( B \) and \( C \) are matrices of the same size as \( A \) such that \( AB = AC \), then \( B = C \). [Hint: Consider \( AB - AC = 0 \).]

2. Give a direct proof of the fact that \( (d) \implies (c) \) in the Invertible Matrix Theorem.

3. Give a direct proof of the fact that \( (c) \implies (b) \) in the Invertible Matrix Theorem.

4. Use the equivalence of \( (a) \) and \( (e) \) in the Invertible Matrix Theorem to prove that if \( A \) and \( B \) are invertible \( n \times n \) matrices, then so is \( AB \).

5. Use the equivalence of \( (a) \) and \( (c) \) in the Invertible Matrix Theorem to prove that if \( A \) and \( B \) are invertible \( n \times n \) matrices, then so is \( AB \).

2.9 Chapter Review

In this chapter we have investigated linear systems of equations. Matrices provide a convenient mathematical representation for linear systems, and whether or not a linear system has a solution (and if so, how many) can be determined entirely from the matrix for the linear system.

An \( m \times n \) matrix \( A = [a_{ij}] \) is a rectangular array of numbers arranged in \( m \) rows and \( n \) columns. The entry in the \( i \)th row and \( j \)th column is written \( a_{ij} \). More generally, such an array, whose entries are allowed to depend on an indeterminate \( t \), is known as a matrix function. Matrix functions can be used to formulate systems of differential equations.

If \( m = n \), the matrix (or matrix function) is called a square matrix.

Concepts Related to Square Matrices

- **Main diagonal**: the entries \( a_{11}, a_{22}, \ldots, a_{nn} \) in the matrix.
- **Trace**: the sum of the entries on the main diagonal.
- **Upper triangular matrix**: \( a_{ij} = 0 \) for \( i > j \).
- **Lower triangular matrix**: \( a_{ij} = 0 \) for \( i < j \).
- **Diagonal matrix**: \( a_{ij} = 0 \) for \( i \neq j \).
- **Transpose**: applying to any \( m \times n \) matrix \( A \), this is the \( n \times m \) matrix \( A^T \) obtained from \( A \) by interchanging its rows and columns.
- **Symmetric matrix**: \( A^T = A \); that is, \( a_{ij} = a_{ji} \).
- **Skew-symmetric matrix**: \( A^T = -A \); that is, \( a_{ij} = -a_{ji} \). In particular, \( a_{ii} = 0 \) for each \( i \).

Matrix Algebra

Given two matrices \( A \) and \( B \) of the same size \( m \times n \), we can perform the following operations:

- **Addition/subtraction**: \( A \pm B \); add/subtract the corresponding elements of \( A \) and \( B \).
- **Scalar multiplication**: \( rA \); multiply each entry of \( A \) by the real (or complex) scalar \( r \).

If \( A \) is \( m \times n \) and \( B \) is \( n \times p \), we can form their product \( AB \), which is an \( m \times p \) matrix whose \((i, j)\)-entry is computed by taking the dot product of the \( i \)th row vector of \( A \) with the \( j \)th column vector of \( B \). Note that, in general, \( AB \neq BA \).
Linear Systems

The general \( m \times n \) system of linear equations is of the form

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\
    \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{align*}
\]

If each \( b_i = 0 \), the system is called homogeneous. There are two useful ways to formulate the above linear system:

1. Augmented matrix:

\[
A^\# = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
    a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
\end{bmatrix}
\]

2. Vector form:

\[
Ax = b,
\]

where

\[
A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}, \quad x = \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}, \quad b = \begin{bmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_m
\end{bmatrix}
\]

Elementary Row Operations and Row Echelon Form

There are three types of elementary row operations on a matrix \( A \):

1. \( P_i \): Permute the \( i \)th and \( j \)th rows in \( A \).
2. \( M_i(k) \): Multiply the entries in the \( i \)th row of \( A \) by the nonzero scalar \( k \).
3. \( A_{ij}(k) \): Add to the elements of the \( j \)th row of \( A \) the scalar \( k \) times the corresponding elements of the \( i \)th row of \( A \).

By performing elementary row operations on the augmented matrix above, we can determine solutions, if any, to the linear system. The strategy is to apply elementary row operations in such a way that \( A \) is transformed into row-echelon form—a process known as Gaussian elimination. By applying back substitution to the linear system corresponding to the row-echelon form obtained, we find the solution. This solution agrees with the solution to the original linear system. If necessary, free parameters may be used to express this solution. A leading one in the far right-hand column of the row-echelon form indicates that the system has no solution.

A row-echelon form matrix is one in which

- All rows consisting entirely of zeros are placed at the bottom of the matrix.
- All other rows begin with a (leading) “1”, called a pivot.
- The leading ones occur in columns strictly to the right of the leading ones in the rows above.
2.9 Chapter Review

Invertible Matrices

An \( n \times n \) matrix \( A \) is invertible if there exists an \( n \times n \) matrix \( B \) such that \( AB = I_n = BA \), where \( I_n \) is the \( n \times n \) identity matrix (ones on the main diagonal, zeros elsewhere). We write \( A^{-1} \) for the (unique) inverse \( B \) of \( A \). One procedure for determining \( A^{-1} \), if it exists, is the Gauss-Jordan technique:

\[
\begin{bmatrix} A | I_n \end{bmatrix} \sim \text{ERO} \cdots \sim \begin{bmatrix} I_n | A^{-1} \end{bmatrix}.
\]

Invertible matrices \( A \) share all of the following equivalent properties:

- \( A \) can be reduced to \( I_n \) via a sequence of elementary row operations.
- The linear system \( Ax = b \) has a unique solution \( x \).
- The linear system \( Ax = 0 \) has only the trivial solution \( x = 0 \).
- \( A \) can be expressed as a product of elementary matrices that are obtained from the identity matrix by applying exactly one elementary row operation.

Additional Problems

Let

\[
A = \begin{bmatrix}
-2 & 4 & 2 & 6 \\
-1 & -1 & 5 & 0
\end{bmatrix}, \quad B = \begin{bmatrix}
-3 & 0 \\
2 & 2 \\
1 & -3 \\
0 & 1
\end{bmatrix}, \quad C = \begin{bmatrix}
-5 \\
-6 \\
3 \\
1
\end{bmatrix}
\]

and \( r = -4 \). For Problems 1–6, compute the given expression, if possible.

1. \( rA - BT \).
2. \( AB \) and \( \text{tr}(AB) \).
3. \( (AC)(AC)^T \).
4. \( (rB)A \).
5. \( (AB)^{-1} \).
6. \( C^T C \) and \( \text{tr}(C^T C) \).

Let

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
2 & 5 & 7
\end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix}
3 & b \\
-4 & a \\
a & b
\end{bmatrix}.
\]

(a) Compute \( AB \) and determine the values of \( a \) and \( b \) such that \( AB = I_2 \).

(b) Using the values of \( a \) and \( b \) obtained in (a), compute \( BA \).

8. Let \( A \) be an \( m \times n \) matrix and let \( B \) be an \( p \times n \) matrix. Use the index form of the matrix product to prove that \( (AB^T)^T = BA^T \).

9. Let \( A \) be an \( n \times n \) matrix.

(a) Use the index form of the matrix product to write the \( ij \)th element of \( A^2 \).

(b) In the case when \( A \) is a symmetric matrix, show that \( A^2 \) is also symmetric.

10. Let \( A \) and \( B \) be \( n \times n \) matrices. If \( A \) is skew-symmetric, use properties of the transpose to establish that \( B^T AB \) is also skew-symmetric.

An \( n \times n \) matrix \( A \) is called nilpotent if \( A^p = 0 \) for some positive integer \( p \). For Problems 11–12, show that the given matrix is nilpotent.

11. \( A = \begin{bmatrix}
3 & 9 \\
-1 & -3
\end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix}
0 & 1 & 1 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}.

12. \( A = \begin{bmatrix}
1 & 1 & 1 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix}
0 & 1 & 1 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}.

\text{\textnormal{(b)}} \quad \text{\textnormal{Using the values of } a \text{ and } b \text{ obtained in (a), compute } BA.}
For Problems 13–16, let
\[
A(t) = \begin{bmatrix}
\frac{e^{-3t} - \sec^2 t}{2} \\
6 \ln t + 3b - 3t
\end{bmatrix}
\]
and
\[
B(t) = \begin{bmatrix}
-7 + t^2 \\
6 - t + 3t^3 + 6t^2 \\
\frac{t}{e^t} - 1 - t^3
\end{bmatrix}.
\]
Compute the given expression, if possible.

13. \(A(t)\).
14. \(\int_0^1 B(t) \, dt\).
15. \(t^3 - A(t) - \sin t - B(t)\).
16. \(B'(t) - e^t A(t)\).

For Problems 17–23, determine the solution set to the given linear system of equations.

17. \(x_1 + 5x_2 + 2x_3 = -6, \quad 4x_2 - 7x_3 = 2, \quad 5x_4 = 0\)
18. \(-5x_1 - x_2 + 2x_3 = 7, \quad -2x_1 + 6x_2 + 9x_3 = 0, \quad -7x_1 + 5x_2 - 3x_3 = -7\)
19. \(x + 2y + z = 1, \quad x + z = 5, \quad 4x + 4y = 12\)
20. \(-x_1 - 2x_2 - x_3 + 3x_4 = 0, \quad -2x_1 + 4x_2 + 5x_3 - 3x_4 = 3, \quad 3x_1 - 6x_2 - 6x_3 + 8x_4 = 2\)
21. \(3x_1 - x_2 + 2x_4 - x_5 = 1, \quad x_1 + 3x_2 + x_3 - 3x_4 + 2x_5 = -1, \quad 4x_1 - 2x_2 + 3x_3 + 6x_4 = -5, \quad x_4 + 4x_5 = -2\)
22. \(x_1 + x_2 + x_3 + x_4 - 3x_5 = 6, \quad x_1 + x_2 + x_3 + 2x_4 - 5x_5 = 8, \quad 2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 = 17, \quad 2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 = 14\)
23. \(x_1 - 3x_2 + 2x_3 = 1, \quad -2x_1 + 6x_2 + 2x_3 = -2\)

For Problems 24–27, determine all values of \(k\) for which the given linear system has (a) no solution, (b) a unique solution, and (c) infinitely many solutions.

24. \(x_1 - kx_2 = 6, \quad 2x_1 + 3x_2 = k\)
25. \(kx_1 + 2x_2 - x_3 = 2, \quad kx_2 + x_3 = 2\)
26. \(kx_1 + x_2 - x_3 = 0, \quad 2x_1 + x_2 - x_3 = 0\)
27. \(x_1 - kx_2 + k^2x_3 = 0, \quad x_1 + kx_1 = 0, \quad x_2 - x_1 = 1\)
28. Do the three planes \(x_1 + 2x_2 + x_3 = 4, x_3 - x_1 = 1, \) and \(x_1 + 3x_2 = 0\) have at least one common point of intersection? Explain.

For Problems 29–34, find a row-echelon form of the given matrix \(A\), determine rank \(A\), and use the Gauss-Jordan technique to determine the inverse of \(A\), if it exists.

29. \(A = \begin{bmatrix} 4 & 7 & 0 \\ -2 & 5 & 3 \end{bmatrix}\)
30. \(A = \begin{bmatrix} 2 & -7 \\ -4 & 14 \end{bmatrix}\)
31. \(A = \begin{bmatrix} 3 & -1 & 6 \\ 0 & 2 & 1 \\ 3 & -5 & 6 \end{bmatrix}\)
32. \(A = \begin{bmatrix} 2 & 0 & 6 \\ 1 & 2 & 0 \\ 0 & 0 & 3 & 4 \end{bmatrix}\)
33. \(A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}\)
34. \(A = \begin{bmatrix} -2 & 3 & 1 \\ 1 & 4 & 2 \\ 0 & 5 & 3 \end{bmatrix}\)

35. Let
\[
A = \begin{bmatrix} 1 & -1 & 3 \\ 4 & -3 & 13 \\ 1 & 1 & 4 \end{bmatrix}.
\]
Solve each of the systems
\[
Ax = e_i, \quad i = 1, 2, 3
\]
where \(e_i\) denote the column vectors of the identity matrix \(I_3\).
36. Solve each of the systems $Ax = b$, if

\[
A = \begin{bmatrix} 2 & 5 \\ 7 & -2 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},
\]

\[
b_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad b_3 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}.
\]

37. Let $A$ and $B$ be invertible matrices.

(a) By computing an appropriate matrix product, verify that $(A^{-1}B)^{-1} = B^{-1}A$.

(b) Use properties of the inverse to derive $(A^{-1}B)^{-1} = B^{-1}A$.

38. Let $S$ be an invertible $n \times n$ matrix and let $k$ be a nonnegative integer. If $A = SDS^{-1}$, prove that $A^k = SDS^k S^{-1}$.

For Problems 39–42, (a) express the given matrix as a product of elementary matrices, and (b) determine the LU decomposition of the matrix.

39. The matrix in Problem 29.
40. The matrix in Problem 32.
41. The matrix in Problem 33.
42. The matrix in Problem 34.

39. (a) Prove that if $A$ and $B$ are $n \times n$ matrices, then

\[
(A + B)^3 = A^3 + 3A^2B + 3AB^2 + B^3.
\]

(b) How does the formula change for $(A - B)^3$?

(c) Can you make a conjecture about the number of terms in the expansion of $(A + B)^k$, in terms of $k$?

44. Suppose that $A$ and $B$ are invertible matrices. Prove that the block matrix

\[
\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}
\]

is invertible.

45. In many different positions can two leading ones of a row-echelon form of a $2 \times 4$ matrix occur? How about three leading ones for a $3 \times 4$ matrix? How about four leading ones for a $4 \times 6$ matrix? How about $m$ leading ones for an $m \times n$ matrix with $m \leq n$?

46. If the inverse of $A^2$ is the matrix $B$, what is the inverse of the matrix $A^{10}$? Prove your answer.

Project: Circles and Spheres via Gaussian Elimination

Part 1: Circles

In this part, we shall see that any three noncollinear points in the plane can be found on a unique circle, and we will use Gaussian elimination to find the center and radius of this circle.

(a) Show geometrically that three noncollinear points in the plane must lie on a unique circle. [Hint: The radius must lie on the line that passes through the midpoint of two of the three points and that is perpendicular to the segment connecting the two points.]

(b) A circle in the plane has an equation that can be given in the form

\[
(x - a)^2 + (y - b)^2 = r^2,
\]

where $(a, b)$ is the center and $r$ is the radius. By expanding the formula, we may write the equation of the circle in the form

\[
x^2 + y^2 + cx + dy = k,
\]

for constants $c, d,$ and $k.$ Using this latter formula together with Gaussian elimination, determine $c, d,$ and $k$ for each set of points below. Then solve for $(a, b)$ and $r$ to write the equation of the circle.

(i) $(2, -1), (3, 3), (4, -1).$
(ii) $(-1, 0), (1, 2), (2, 2).$
Part 2: Spheres  In this part, we shall extend the ideas of Part 1 and consider four noncoplanar points in 3-space. Any three of these four points lie in a plane but are noncollinear (why?). A sphere in 3-space has an equation that can be given in the form

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2,$$

where $(a, b, c)$ is the center and $r$ is the radius. By expanding the formula, we may write the equation of the sphere in the form

$$x^2 + y^2 + z^2 + ux + vy + wz = k,$$

for constants $u, v, w$, and $k$.

(a) Using the latter formula above together with Gaussian elimination, determine $u, v, w$, and $k$ for each set of points below. Then solve for $(a, b, c)$ and $r$ to write the equation of the sphere.

(i) $(1, -1, 2), (2, -1, 4), (-1, -1, -1), (1, 4, 1)$.
(ii) $(2, 0, 0), (0, 3, 0), (0, 0, 4), (0, 0, 6)$.

(b) What goes wrong with the procedure in (a) if the points lie on a single plane? Choose four points of your own and carry out the procedure in part (a) to see what happens. Can you describe circumstances under which the four coplanar points will lie on a sphere?